JOURNAL OF VIROLOGY, Jan. 2004, p. 1032–1038 Vol. 78, No. 2
0022-538X/04/$08.00+0 DOI: 10.1128/JVI.78.2.1032–1038.2004

High Fidelity of Yellow Fever Virus RNA Polymerase

Konstantin V. Pugachev, Farshad Guirakhoo, Simeon W. Ocran, Fred Mitchell, Megan Parsons, Caroline Penal, Soheila Girakhoo, Svetlana O. Pougatcheva, Juan Arroyo,† Dennis W. Trent, and Thomas P. Monath*

Acambis, Inc., Cambridge, Massachusetts 02139

Received 4 June 2003/Accepted 3 October 2003

Three consecutive plaque purifications of four chimeric yellow fever virus-dengue virus (ChimeriVax-DEN) vaccine candidates against dengue virus types 1 to 4 were performed. The genome of each candidate was sequenced by the consensus approach after plaque purification and additional passages in cell culture. Our data suggest that the nucleotide sequence error rate for SP6 RNA polymerase used in the in vitro transcription step to initiate virus replication was as high as 1.34×10^{-4} per copied nucleotide and that the error rate of the yellow fever virus RNA polymerase employed by the chimeras for genome replication in infected cells was as low as 1.9×10^{-7} to 2.3×10^{-7}. Clustering of beneficial mutations that accumulated after multiple virus passages suggests that the N-terminal part of the prM protein, a specific site in the middle of the E protein, and the NS4B protein may be essential for nucleocapsid-envelope interaction during flavivirus assembly.

A fundamental feature of RNA viruses is the error-prone nature of replication of their genomes, resulting in adaptation and rapid evolution. Various in vitro, in vivo, and ex vivo methods have been used to estimate mutation rates of viral RNA-dependent RNA polymerases. The generally accepted average value of 10^{-4} to 10^{-5} mutations per nucleotide (nt) per round of RNA replication (reviewed in references 6, 7, 23, and 28) results in substitution of 0.1 to 1 nt per each synthesized genomic RNA molecule of ~10 kb. Mutation rates vary between 10^{-3} and 10^{-6} in individual studies, depending on the virus and method used, although none of these methods provide an exact measure of the error rate. In vitro methods that utilize purified viral RNA polymerases may overestimate the error rate due to suboptimal in vitro enzymatic conditions and experimental complexity. In ex vivo methods, such as sequencing of phenotypic mutants generated in cell culture and analysis of monoclonal antibody escape mutants or of rates of reversion of phenotypic markers, the number of rounds of RNA replication in infected cells is not precisely known. Methods that involve a reverse transcription-PCR step with cloning of cDNA in bacteria prior to sequencing can be misleading because of the accumulation of additional mutations introduced by reverse transcription-PCR.

The *Flavivirus* genus of the *Flaviviridae* family consists of about 70 viruses of which 40 have been associated with human illness (reviewed in reference 2). Flaviviruses are small enveloped plus-strand RNA viruses (reviewed in reference 18). The viral particle contains a nucleocapsid composed of viral RNA and capsid protein C, surrounded by a lipid bilayer in which the envelope protein E and membrane protein M are embedded. The genomic RNA is about 11,000 nt long and contains a single long open reading frame (ORF). Translation of the ORF produces a single polyprotein precursor containing viral proteins in the order: C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-2K-NS4B-NS5, where C through E are the structural components of the virion (prM is a precursor for M). Nonstructural (NS) proteins NS1 through NS5 are required for virus replication in the cytoplasm of infected cells (18). Immature virions are produced by budding into the lumen of the endoplasmic reticulum (ER) and movement of virus to the cell surface through the exocytosis pathway. The structure of the flavivirus particle has been resolved by cryoelectron microscopy and fitting the known structure of the E protein (29) into the electron density map (15).

The ChimeriVax technology has been used to create attenuated flavivirus vaccines using the yellow fever virus (YF) 17D vaccine virus as a vector in which the prM-E envelope protein genes are replaced with those from a heterologous flavivirus (4, 10, 11, 24–26). To produce a chimeric virus, the viral RNA genome is first synthesized by in vitro transcription of an appropriately engineered DNA template with bacteriophage SP6 RNA polymerase. Following transfection of cells with the in vitro RNA transcripts, virus replication begins with genome amplification by the YF 17D RNA-dependent RNA polymerase (YFpol). In this study, yellow fever virus-dengue virus (ChimeriVax-DEN) type 1 to 4 chimeras were prepared by transfecting Vero cells with in vitro RNA transcripts. The resulting viruses were cloned by three consecutive plaque purifications. Both cloned and uncloned viruses were sequentially passaged to evaluate their genetic stability during prolonged growth in cell culture. Each virus was sequenced by the consensus method several times at different passages, including before and after plaque purification of the cloned variants. We observed that YFpol introduced very few mistakes in the genome during plaque purification steps, indicating a high fidelity of the polymerase. Characterization of genomes with YFpol mistakes and mutations that accumulated during multiple passages shed new light on the mechanism of flavivirus evolution, RNA replication, and particle assembly. Data from these studies challenge the widely held concept that RNA genomes accumulate substitution of up to 1 nt for each synthesized genomic RNA molecule.

* Corresponding author. Mailing address: Acambis, Inc., 38 Sidney St., Cambridge, MA 02139. Phone: (617) 761-4200. Fax: (617) 494-1741. E-mail: tom.monath@acambis.com.
† Present address: DynPort Vaccine Co., LLC, Frederick, MD 21702.

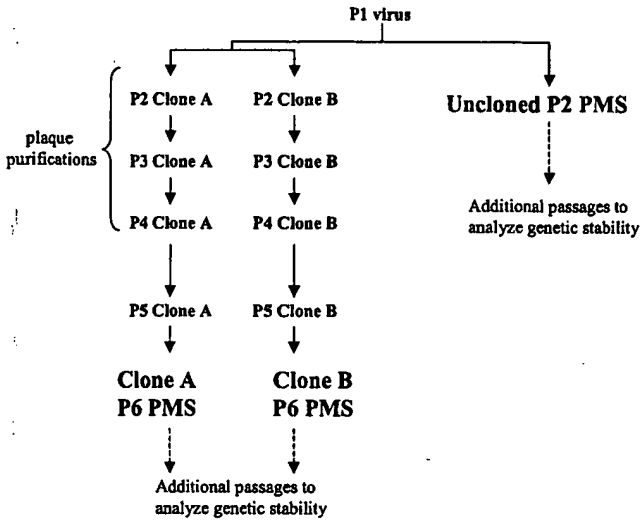

FIG. 1. Passages of the cloned and uncloned ChimeriVax-DEN2$_{2000}$ viruses. P1 virus was obtained by electroporation of Vero cells with in vitro RNA transcripts synthesized using SP6 RNA polymerase. Uncloned P2 virus was generated by one additional passage. Cloned PMS candidates (clones A and B) were produced by three consecutive plaque purifications starting from P1 virus followed by two additional passages. Genetic stabilities of these cloned and uncloned variants were examined in additional downstream virus passages. ChimeriVax-DEN2$_{2001}$, -DEN3, -DEN4, and -DEN1 viruses were passaged similarly, except that plaque purification passages (P3 to P5) were initiated from the uncloned P2 viruses.

Virus passages. A diagram of passages used to prepare uncloned passage 2 (P2) and two cloned P6 Pre-Master Seed (PMS) variants of ChimeriVax-DEN2$_{2000}$ virus is shown in Fig. 1. Uncloned and cloned ChimeriVax-DEN2$_{2001}$, -DEN3, -DEN4, and -DEN1 PMS viruses were prepared similarly, except that plaque purification passages (P3 to P5) were initiated at the uncloned P2 level rather than P1. Vero cells provided by Aventis Pasteur (Lyon, France) were used in all experiments. Cells and viruses were grown in Earl's minimum essential medium (EMEM) supplemented with L-glutamine, MEM nonessential amino acids (Gibco), and 10% fetal bovine serum (FBS; HyClone). Medium 199 (Gibco) containing 3% FBS was used for plaque purifications done in six-well culture plates by the agarose double overlay method with neutral red staining described previously (24).

In vitro RNA transcripts were synthesized with SP6 RNA polymerase (Epicentre) on full-length DNA templates for the four ChimeriVax-DEN viruses prepared using two- or three-fragment ligations (11). No PCR steps were involved in preparation of the templates, and each plasmid was grown from a single bacterial colony. The P1 viruses were obtained by electroporation of 1.5×10^7 cells (which had been trypsinized and washed with phosphate-buffered saline) with 50 to 200 ng of the in vitro RNA transcript, followed by seeding of the cells into T75 or T25 flasks. P1 virus was harvested on days 3 to 5 posttransfection. The uncloned P2 PMS ChimeriVax-DEN1 to -DEN4 viruses, cloned P6 ChimeriVax-DEN2$_{2000}$ virus, and cloned P7 ChimeriVax-DEN1, -DEN3, -DEN4, and -DEN2$_{2001}$ PMS viruses were obtained by infection of cells at a controlled multiplicity of infection (MOI) of 0.001 PFU/cell. Titers of these virus samples were in excess of 10^6 PFU/ml. Up to 13 additional (genetic stability) passages of these viruses were done without intermediate virus titrations at an assumed MOI of 0.001 PFU/cell. All virus harvests were collected when ~5 to 10% cytopathic effect was observed. Virus-containing media were clarified by low-speed centrifugation, diluted with an equal volume of FBS, and stored at −70°C. Up to 10 clones were prepared for each chimera by three sequential plaque purifications. Plaque picks (agarose plugs) were placed in 0.4 ml of medium 199 (Gibco) containing 50% FBS, vortexed for several seconds, and frozen at −70°C. To sequence viruses obtained after the first plaque purification, 100 μl of the indicated plaque pick supernatant (virus eluted from agarose plug) was used to infect cells in T25 flasks, and the amplified viruses were harvested and sequenced. Otherwise, all samples of virus passages shown in Table 1 were sequenced without additional expansion.

Sequence analysis. Consensus sequencing of virus samples was performed as previously described (27a). Sequence data were aligned and analyzed using Sequencher 3.0 (GeneCodes) software. Nucleotide heterogeneities were registered when a heterogeneous signal was observed in all chromatograms representing both plus and minus cDNA strands. For some viruses (Table 1), only the first of the five cDNA amplicons (fragment I) was prepared and sequenced, including the 5′ untranslated region (5′ UTR) and the structural protein genes. In several cases, such as viruses obtained after the first plaque purification or some intermediate genetic stability passages, only the regions where mutations were detected at later passages were sequenced, assuming that no other mutations were present in the rest of the genome (or fragment I).

Consensus full-genome sequencing, which does not detect random mutations, of the uncloned P2 viruses revealed no unexpected mutations (Table 1). Mutations accumulated in the uncloned ChimeriVax-DEN1 to -DEN3 variants during their further passages to study genetic stability (see passages after P2 in Table 1).

Since no mutations were detected in the uncloned P2 ChimeriVax-DEN2$_{2000}$ virus, no mutations should be detected in the consensus sequence of P1 virus used to start plaque purifications. Nevertheless, each of the cloned A and B P6 ChimeriVax-DEN2$_{2000}$ viruses had a 1-nt change: G to T at nt 7734, resulting in an R-to-M amino acid change at NS5-29 located in a relatively conserved N-terminal region of NS5 and a silent A-to-T change at nt 6643, respectively (Table 1). Two additional genetic stability passages of clone A virus resulted in accumulation of subpopulations containing changes at nt 515 and 2135 detected at P8. Changes in clone A had potentially significant phenotypic implications. Therefore, further work with this clone was terminated. In clone B virus at P9, an additional G/C heterogeneity was observed at nt 517. The G-to-C nt change at this position resulted in an M-to-I amino acid change at prM-12 in the N-terminal part of prM that is cleaved during virus maturation. The effect of this mutation is similar to that of the nt 515 change in clone A, resulting in a similar mutation (M to L) at the same residue of prM. An M-to-L change at this residue was also seen in uncloned P15 ChimeriVax-DEN2$_{2001}$ virus (Table 1). This heterogeneity became a dominating change in virus at P14. An L-to-F change at amino acid 177 of NS4B protein was selected during virus

TABLE 1. Nucleotide and amino acid changes in uncloned and cloned ChimeriVax-DEN viruses

Virus	Version	Mutations detected at sequenced passages[a]
ChimeriVax-DEN2$_{2000}$	Uncloned	P2, none[b]; P5, none[b]; P10, $A_{1716}T$ ($K_{E\text{-}246}M$), $G_{3287}A$ ($V_{NS1\text{-}275}M$)[b]
	Clone A	**P2, $G_{7734}T$ ($R_{NS5\text{-}29}M$)[c]; P6, $G_{7734}T$ ($R_{NS5\text{-}29}M$)[b]**; P8, A_{515}/C ($M_{prM\text{-}12}/L$), C_{2135}/G ($Q_{E\text{-}386}/E$), $G_{7734}T$ ($R_{NS5\text{-}29}M$)[b]
	Clone B	**P2, $A_{6643}T$ (silent)[c]; P6, $A_{6643}T$ (silent)[b]**; P9, G_{517}/C ($M_{prM\text{-}12}/I$), $A_{6643}T$ (silent)[c]; P14, $G_{517}C$ ($M_{prM\text{-}12}I$), $A_{6643}T$ (silent), $C_{7427}T$ ($L_{NS4B\text{-}177}F$)[c]; P18, $G_{517}C$ ($M_{prM\text{-}12}I$), $A_{6643}T$ (silent), $C_{7427}T$ ($L_{NS4B\text{-}177}F$)[b]
ChimeriVax-DEN2$_{2001}$	Uncloned	P2, none[b]; P15, $A_{515}C$ ($M_{prM\text{-}12}L$), $G_{1321}A$ (silent), $A_{1590}G$ ($K_{E\text{-}204}R$), $A_{4046}T$ ($I_{NS2A\text{-}176}L$), $T_{5485}C$ (silent), $C_{6888}T$ ($A_{2K\text{-}20}V$), $C_{7427}T$ ($L_{NS4B\text{-}177}F$), G_{9094}/T (silent), $C_{10307}T$ (silent)[b]
	Clone A	**P6, none[b]**; P7, none[b]; P10, $T_{551}G$ ($L_{prM\text{-}24}V$), $G_{1730}T$ ($V_{E\text{-}251}F$)[b]; P20, $T_{551}G$ ($L_{prM\text{-}24}V$), $G_{1730}T$ ($V_{E\text{-}251}F$)[b]
	Clone B	**P3, none[c]; P6, $A_{1030}G$ (silent), $C_{1789}T$ (silent)[d]**; P15, $A_{1030}G$ (silent), $A_{1590}G$ ($K_{E\text{-}204}R$), $C_{1789}T$ (silent), $T_{2444}C$ (silent), $C_{7443}T$ ($A_{NS4B\text{-}182}V$), $A_{8677}G$ (silent)[b]
ChimeriVax-DEN3	Uncloned	P2, none[b]; P5, none[c]; P10, $C_{564}A$ ($A_{prM\text{-}28}D$), $T_{2561}C$ ($Y_{NS1\text{-}35}H$)[c]; P15, $C_{564}A$ ($A_{prM\text{-}28}D$), $T_{2561}C$ ($Y_{NS1\text{-}35}H$), A_{6502}/G (silent), A_{7231}/G ($I_{NS4B\text{-}113}/M$)[b]
	Clone A	**P3, $C_{6607}T$ (silent)[c]; P6, $C_{6607}T$ (silent)[b]**; P7, $C_{6607}T$ (silent)[b]; P10, $C_{6607}T$ (silent)[c]; P15, $C_{6607}T$ (silent), C_{7859}/T (silent)[c,e]; P20, $C_{6607}T$ (silent), $C_{7421}T$ ($L_{NS4B\text{-}177}F$), C_{7859}/T (silent)[b]
ChimeriVax-DEN4	Uncloned	P2, none[b]; P15, none[b]
	Clone B	**P6, none[b]**; P7, none[b]; P10, none[b]; P20, none[b]
ChimeriVax-DEN1	Uncloned	P2, none[b]; P5, A_{1590}/G ($K_{E\text{-}204}/R$), G_{1730}/T ($V_{E\text{-}251}/F$)[c]; P15, $A_{1590}G$ ($K_{E\text{-}204}R$), G_{1730}/T ($V_{E\text{-}251}/F$), A_{7237}/G ($I_{NS4B\text{-}113}/M$), C_{7466}/T ($P_{NS4B\text{-}190}/S$)[b]
	Clone A	**P3, $G_{1730}T$ ($V_{E\text{-}251}F$), $C_{2282}A$ ($L_{E\text{-}435}I$)[c]; P7, $G_{1730}T$ ($V_{E\text{-}251}F$), $C_{2282}A$ ($L_{E\text{-}435}I$)[b]**
	Clone B, D	**P3, $G_{1730}T$ ($V_{E\text{-}251}F$)[c]; P7 or P6, $G_{1730}T$ ($V_{E\text{-}251}F$)[b]**
	Clone C	**P3, $G_{1912}T$ ($E_{E\text{-}311}D$), $G_{2030}T$ ($V_{E\text{-}351}L$)[c]; P6, $G_{1912}T$ ($E_{E\text{-}311}D$), $G_{2030}T$ ($V_{E\text{-}351}L$)[b]**
	Clone E	**P3, $A_{1590}G$ ($K_{E\text{-}204}R$), $A_{3952}T$ (silent)[c]; P6, $A_{1590}G$ ($K_{E\text{-}204}R$), $A_{3952}T$ (silent)[b]**
	Clone F	**P3, $C_{788}T$ (silent), $A_{1590}G$ ($K_{E\text{-}204}R$)[d]**
	Clone G	**P3, $G_{1730}T$ ($V_{E\text{-}251}F$)[d]**
	Clone H	**P3, $G_{1912}T$ ($E_{E\text{-}311}D$), $G_{2030}T$ ($V_{E\text{-}351}L$)[d]**
	Clone I	**P3, $A_{1590}G$ ($K_{E\text{-}204}R$)[d]**
	Clone J	**P3, none[b]; P6, none[b]**; P7, none[b]; P10, $A_{1590}G$ ($K_{E\text{-}204}R$)[b]; P20, $A_{1590}G$ ($K_{E\text{-}204}R$), G_{6966}/T ($S_{NS4B\text{-}23}/I$), G_{7190}/A ($V_{NS4B\text{-}98}/I$)[b]

[a] The passage number is shown first; any mutation(s) detected follow the passage number. Nucleotide changes at indicated positions (numbering from the beginning of the genome) and amino acid changes (in parentheses) at indicated protein residues [from the N terminus of each protein; amino acid numbering of YF proteins as in references 3 and 17] are shown. A heterogeneous signal is indicated by a slash. Data for clones after the first and/or last plaque purifications are in bold type.

[b] Full genome sequenced.

[c] Sites of mutations found after last plaque purification (in samples after first plaque purification) or at a later genetic stability passage were sequenced only. It can be assumed that the full genome (or fragment I) sequence is known.

[d] Only fragment I sequenced, including the entire structural protein coding region.

[e] In independent passages of the DEN3 clone A virus to P15 from a Master Seed (large-scale manufacturing P8), accumulation of $A_{1584}G$ ($K_{E\text{-}202}R$; identical to the E-204 mutation in the DEN1 and DEN2 chimeras) and $C_{7437}T$ ($A_{NS4B\text{-}182}V$) mutations was documented, in addition to the silent nt 6607 change (F. Mitchell, data not shown).

passage to P14. The P18 of clone B had a consensus sequence identical to that of P14. It was significant that the P2 samples of both clones A and B harvested after the first plaque purification contained the same mutations as the P6 viruses at nt 7734 and 6643, respectively (Table 1).

Starting plaque purification with the uncloned P2 virus, rather than with P1 virus, immediately yielded acceptable cloned vaccine candidates of ChimeriVax-DEN2$_{2001}$ (clone A), -DEN3 (clone A), and -DEN4 (clone B; clone A was not sequenced) (Table 1). Only clone A of the ChimeriVax-DEN3 chimera had one silent nucleotide change at position 6607 detected at both P3 after the first plaque purification and P6 after the third plaque purification. In contrast, 10 clones of ChimeriVax-DEN1 were examined to find one without unacceptable mutations. Only clone J of the ChimeriVax-DEN1 chimera had no mutations at both P3 and P6. The entire genomes of clones A, B, C, and E were sequenced at P6, and the presence of mutations was checked at P3. Only the structural protein coding region of clone D was sequenced at both P3 and P6. In all cases, mutations present at P6 were present in P3 viruses (Table 1). (Fragment I of clones F to I was sequenced only at P3.)

Some clones of the ChimeriVax-DEN1 chimera had identical mutations, such as the E-251 V-to-F change in clones A, B, D, and G; the E-204 K-to-R change in clones E, F, and I; and the E-311 E-to-D and E-351 V-to-L changes in clones C and H. These changes were not detected in the consensus sequence of the starting uncloned P2 virus. Mutations E-204 and E-251 were first detected as heterogeneities in the uncloned P5 virus. The E-204 mutation became dominant in the uncloned P15 virus, and it also appeared in clone J virus by P10 (Table 1). The rate of accumulation of the E-204 mutation was studied in more detail in additional passaging experiments using clone J. These studies indicated this mutation was infrequently introduced by YFpol (C. Penal and F. Mitchell, data not shown). Since the E-204 mutation and all other mutations found in clones A to I at P3 appeared as clear mutations, they must have been present in individual particles of the uncloned P2 virus population before the first plaque purification. The fact that the E-204, E-251, E-311, and E-351 mutations were not de-

tected in the consensus P2 sequence despite the high proportion of these mutations in the plaques picked was most likely due to the low sensitivity of consensus sequencing. Our prior experience with quantitation of mutant subpopulations using the MAPREC method (5) demonstrated that the level of detection of mutations by sequencing is approximately 10 to 15%. During the first plaque purification step, a few slightly larger plaques were picked together with smaller plaques which could account for some of the observed variation. The possibility that any of these mutations originated in plasmid templates is unlikely because of the high fidelity of bacterial DNA synthesis (7, 23). Ten individual plasmid clones from the ChimeriVax-DEN1 DNA template were prepared in *Escherichia coli* and sequenced. None of the clones contained the mutations detected in clones A to I of the virus or any other mutations (M. Parsons, data not shown).

Clone B of ChimeriVax-DEN2$_{2001}$ was prepared and sequenced at P3 after the first plaque purification, P6 after the third plaque purification, and P15. The P6 virus had two silent nucleotide changes at nt 1030 and 1789 which were not present at P3 (Table 1). These mutations occurred during plaque purification passages, and thus represented mistakes of YFpol. These were the only two mistakes introduced by YFpol during three plaque purifications of all the cloned chimeras described.

The cloned and uncloned ChimeriVax-DEN4 viruses did not accumulate mutations during genetic stability passages. In contrast, both cloned and uncloned variants of the other chimeras accumulated various mutations in genetic stability passages. Most of these mutations resulted in amino acid changes, indicating that virus adaptation is driven by forces operating at the protein function level. Some of the mutations observed among different chimeras coincided (Table 1). For instance, the same E-204 mutation accumulated in uncloned and clone B ChimeriVax-DEN2$_{2001}$ (detected at P15 of both variants), uncloned and clone J of ChimeriVax-DEN1 (at P5 and P10, respectively), and ChimeriVax-DEN3 (the E-202 mutation; see footnote *e* to Table 1). This suggests convergent mechanisms of adaptation of the ChimeriVax-DEN1 to -DEN3 chimeras to Vero cells. Many diverse mutations accumulated in uncloned and cloned versions of the same chimera or in different clones of a chimera (Table 1). The latter observation indicates that one virus can evolve in different directions under similar growth conditions. It is possible that several different mutations can improve virus replication and that further virus evolution depends on which mutation occurs first.

Classification of mutations. The interpretation of mutations observed at different passages of cloned and uncloned ChimeriVax-DEN viruses is summarized in Table 2. No mutations were detected in the consensus sequences of all chimeras in uncloned P2 virus. As a rule, mutations in cloned viruses detected after three consecutive plaque purifications (at P6 or P7), were also observed after the first plaque purification. We believe that these mutations were present in individual virions present in the uncloned P1 or P2 virus population obtained after electroporation of cells with in vitro RNA transcripts. These mutant viruses were randomly selected by the first plaque purification, as no additional mutations occurred during subsequent plaque purification passages. Since each act of plaque formation requires multiple amplifications of the genomic RNA in the initiating viral particle, YFpol must have high fidelity. Each act of plaque formation is a very harsh virus passage in which replication is allowed to proceed until complete destruction of participating cells. Therefore, it is unlikely that mutations were introduced by YFpol in the process of virus replication prior to the first plaque purification. Most likely, they were introduced by SP6 RNA polymerase during in vitro RNA synthesis. This rule applies to all cloned viruses, except for clone B of ChimeriVax-DEN2$_{2001}$ (Table 1). This virus had two silent nucleotide changes at P6 level (nt 1030 and 1789), which were not present at P3. Thus, these two mutations were the only detected mistakes of YFpol that occurred during plaque purifications.

TABLE 2. Assigned classification of mutations detected during preparation of ChimeriVax-DEN viruses and the genetic stability passages shown in Table 1

Type of mutation	Mutation detected			
	In uncloned P2 virus	After the first plaque purification[a]	After the third plaque purification[b]	In genetic stability passages
Introduced by SP6 polymerase (17 nucleotide changes in 11.5 genomes)	−	+	+	+
Introduced by YFpol (2 nucleotide changes in 10.5 genomes)	−	−	+	+
Beneficial mutation[c] (15 and 19 nucleotide changes in the cloned and uncloned variants, respectively)	−	−	−	+

[a] P2 plaque purification step of cloned ChimeriVax-DEN2$_{2000}$ or P3 of cloned ChimeriVax-DEN1, -DEN2$_{2001}$, -DEN3, and -DEN4 (Table 1) (Fig. 1).

[b] P6 of cloned ChimeriVax-DEN2$_{2000}$ or P6 or P7 of cloned variants of other chimeras (Table 1).

[c] Beneficial mutations that accumulated in genetic stability passages of cloned variants were introduced by YFpol, while it is unknown which polymerase (SP6 or YFpol) introduced mutations that accumulated in uncloned variants. The E-204 and E-251 mutations beneficial for ChimeriVax-DEN1 were also found after plaque purifications in some of the clones of the chimera.

All mutations in individual cloned viruses selected by plaque purifications were stably fixed. As a rule, these mutations did not accumulate in other clones or uncloned variants during genetic stability passages. Other mutations accumulated in genetic stability passages of most clones and uncloned variants. Subpopulations containing such mutations outgrew the nonmutant viruses in the population, and therefore such mutations are beneficial. All the chimeras are viable immediately after transfection. However, analyses of some of the accumulating virus variants showed that these variants produce larger plaques in Vero cells and higher peak titers compared to nonmutants (by up to 1 $\log_{10}$ PFU/ml; data not shown). Thus, there is room for improvement of growth of these artificial chimeras.

Some mutations in ChimeriVax-DEN1 clones A to I detected after the first and the third plaque purification (Table 1) did not clearly fit the classification in Table 2. The E-204 and

E-251 mutations were detected in more than one clone and also accumulated in genetic stability passages of the uncloned virus. The E-204 mutation also accumulated during continuous passage of ChimeriVax-DEN1 clone J. Thus, these mutations are beneficial for the ChimeriVax-DEN1 chimera replication. Since mutations are introduced by YFpol infrequently, these changes were likely mistakes of SP6 polymerase amplified prior to the first plaque purification.

Of the 17 nucleotide changes considered to be mistakes of SP6 polymerase, 5 were transitions and 12 were transversions. Thirteen of the mutations were changes from A, G, or C nucleotides to T, with a G-to-T change being the most frequent (hine mutations). This suggests that a misincorporation of T was the most common mistake during in vitro transcription. Of the 17 nucleotide changes introduced by YFpol, including the two mistakes that occurred during plaque purifications plus beneficial mutations that accumulated during genetic stability passages of the clones, 11 changes were transitions and 6 were transversions. The most frequent changes were A to G (four mutations) and C to T (five mutations), which theoretically should be the two most common mistakes during late plus-strand RNA synthesis due to T · G and G · T mispairing, respectively. This observation may point toward the existence of distinct mechanisms for synthesis of double-stranded RNA (with high fidelity), representing most rounds of amplification in infected cells, and late transcription (with low fidelity). The majority of rounds of viral RNA synthesis should occur during the latent phase of infection accompanied by exponential duplication of the double-stranded replicative-form (RF) RNA. During the late linear phase, RNA synthesis is mostly single-round transcription on existing minus-strand template that is stable for at least 20 h and is renewed slowly (reviewed in references 1a and 30).

Estimates of error rates for SP6 RNA polymerase and YFpol. Sequences of 11.5 equivalents of complete virus genome determined after the first plaque purification (sequences of fragment I are counted as 0.25 genome) contain a total of 17 nucleotide changes (Table 1) considered to be mistakes of SP6 polymerase introduced during a single round of DNA-dependent RNA synthesis in vitro. Thus, the error rate of SP6 RNA polymerase is 17/11.5 = 1.48 misincorporations per one genomic RNA molecule synthesized (~11,000 nt long), or 1.34×10^{-4} per nucleotide copied, which is comparable to the reported error rate of T7 RNA polymerase (0.5×10^{-4} [13]). However, it could be somewhat lower because all of the mutations may not be due to mistakes of SP6 polymerase and the accumulation of mutations in ChimeriVax-DEN1 may have occurred prior to the first plaque purification. On the other hand, if plaque purifications of ChimeriVax-DEN2$_{2001}$, -DEN3, and -DEN4 had been done at P1, rather than P2, more mistakes of SP6 polymerase may have been detected.

Two nucleotide changes were present in a total of 10.5 full genome equivalents that occurred during plaque purifications and therefore were clearly introduced by YFpol (Table 1). To estimate the YFpol error rate, we need to know how many times a genomic RNA molecule is copied during one act of plaque formation. Chimeras grow to titers of $\sim 10^7$ PFU/ml in Vero cells; therefore, a plaque pick of roughly 100 μl in volume will contain $\sim 10^6$ infectious particles. We assume that 100 to 1,000 times more plus strands are synthesized (10^8 to 10^9 molecules) and are not accounted for in infectious virus because of the loss of infectivity of mature secreted particles, as well as the accumulation of immature viral particles and possible RNA turnover inside the cells (reviewed in references 1a and 30). If most rounds of RNA amplification are semiconservative duplications of the RF RNA, as many as 25 to 30 rounds of synthesis can take place (e.g., 10^9 molecules $\approx 2^{30}$, thus requiring 30 duplications) during one plaque formation, in $\sim 10^3$ to 10^4 participating cells. In support of our estimate of the total number of plus strands synthesized, 20 to 40 μg of protein E per 10^7 cells is produced following infection with tick-borne encephalitis virus (1, 14a). Since each viral particle contains 180 protein E monomers (15), up to 2×10^5 viral particles (equal to secreted plus strands) are produced by each cell. This number does not include all plus strands synthesized in infected cell.

By dividing the number of YFpol mistakes (two mistakes) by the number of sequenced full genome equivalents (10.5 genomes), the estimated number of rounds of RNA synthesis during one plaque formation (25 to 30 rounds), and the number of plaque purification steps for each clone (three steps), the resulting error rate of YFpol is $2/[10.5 \times (25 \text{ to } 30) \times 3] = 0.0021 - 0.0025$ misincorporation per genomic RNA molecule (11,000 nt long) per synthesis, or 1.9×10^{-7} to 2.3×10^{-7} per copied nucleotide. With our system, we can detect only nonlethal mutations. We do not believe, however, that lethal mutations could constitute a percentage significant enough to affect the estimate, considering the overall low occurrence of mutations during plaque purifications. It is known that the distribution of mutations in RNA viruses is shifted toward neutral or slightly deleterious mutations rather than strongly deleterious or lethal mutations that are quickly eliminated by competition during virus replication; besides, a small fraction of mutations is beneficial (8, 9). To conclude that there could be a large proportion of undetected lethal mutations, we must observe many synonymous mutations, such as those that do not change the amino acid encoded, which is not the case.

Our data indicate that the mutation rate of YFpol is on the low end of the spectrum among RNA polymerases. This observation is in accord with previous studies demonstrating the high genetic stability of YF (12, 27, 33) and makes YF 17D an excellent vector for the construction of recombinant vaccines. It is believed that the evolution of YF and other flaviviruses in nature is constrained by the need of alternation of vertebrate and arthropod hosts. Our data extend this hypothesis further, suggesting that in the course of flavivirus evolution, variants with high fidelity of their RNA polymerases were selected to facilitate the necessary continuous host alternation.

Hot spots for beneficial amino acid changes. All beneficial amino acid changes that accumulated during genetic stability passages of the cloned and uncloned viruses were plotted in Fig. 2. A conspicuous clustering of mutations was observed, specifically in the N-terminal part of prM in ChimeriVax-DEN2 and -DEN3 chimeras, the central part of the E protein in ChimeriVax-DEN1, -DEN2, and -DEN3, and in the NS4B protein in ChimeriVax-DEN1, -DEN2, and -DEN3. One could argue that if we adapted wild-type DEN viruses or YF 17D to Vero cells, the same mutations could be expected. This appears not to be the case. Specific mutations in the YF-specific NS4B protein did not accumulate in the ChimeriVax-DEN4

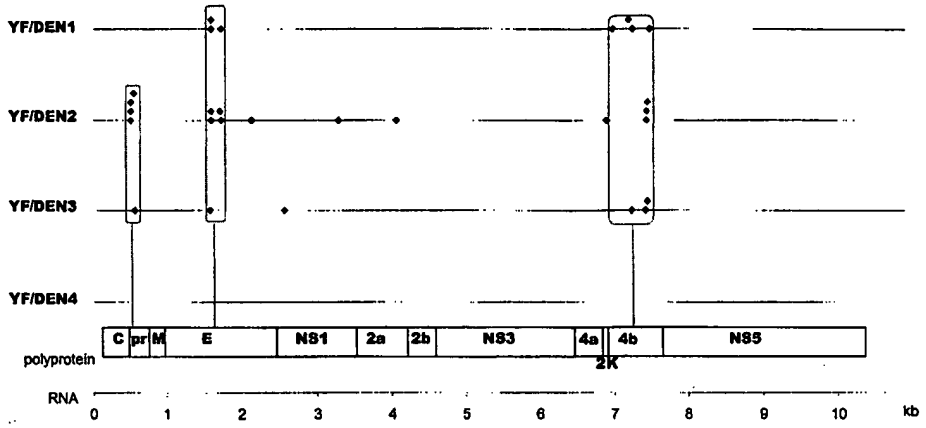

FIG. 2. Clustering of beneficial amino acid changes in the four ChimeriVax-DEN chimeras that accumulated in all cloned and uncloned viruses during genetic stability passages. The locations of mutations in the polyprotein and genomic RNA in each of the chimeras were plotted according to their positions in the genome.

virus. A recent study on adaptation of tick-borne encephalitis virus grown in BHK-21 cells demonstrated that the majority of adaptive mutations resulted in an increase in the net positive charge of the E protein, presumably facilitating the protein interaction with glycosaminoglycans on the cell surface (22). None of the observed mutations in the E protein in ChimeriVax-DEN1 (E-204 K-to-R and E-251 V-to-F mutations), ChimeriVax-DEN2 (E-204 K-to-R, E-246 K-to-M, E-251 V-to-F, and E-386 Q-to-E mutations), and ChimeriVax-DEN3 (E-202 K-to-R mutation) were of this kind. Some of the mutations could affect the ability of the viruses to antagonize cell-specific (e.g., RNA interference) or -nonspecific antiviral mechanisms. It is still unclear why no mutations accumulated in the ChimeriVax-DEN4 chimera.

We favor another possibility—that these mutations are adaptations in response to replacing the YF-specific envelope with the heterologous DEN envelopes. An imperfect match between the YF-specific capsid and the DEN1- to DEN3-specific envelopes dictated adaptation during virus growth. The DEN4 chimera needed no adaptation due to an incidental perfect match between the capsid and DEN4 envelope. Theoretically, this is possible, despite the fact that DEN4 is a type of the same virus species because the specific requirements for interaction between the capsid and envelope are unknown.

If the latter hypothesis is correct, it is interesting that no mutations accumulated in the YF-specific capsid protein C. The mutations in the N-terminal part of prM could have functional significance in light of recent data indicating that prM functions as a chaperone necessary for proper folding of E (20). These mutations could facilitate a more appropriate folding of E such that the envelope more precisely conforms to the architecture of the nucleocapsid. All mutations in the cluster located in the middle of E protein are on the outward surface of the protein in the dimerization domain II (29). The effects of these changes on the structure and function of the E proteins of chimeric viruses need further investigation.

Most intriguing is the cluster of mutations in NS4B (Fig. 2), because it suggests that this protein plays a role in virus assembly. Flavivirus replication induces proliferation of smooth ER membranes and the accumulation of characteristic membranous structures described as convoluted membranes (CM), paracrystalline arrays (PC), and vesicle packets (VP) (21, 31, 32). Viral double-stranded RNA together with NS1, NS3, NS5, NS2A, and NS4A proteins colocalized with VP structures, indicating that they are the main sites of RNA synthesis, while CM and PC structures were proposed to be main sites of proteolytic cleavage (14, 21, 32). NS4B appears to have a distinct distribution, associated with ER membranes throughout the cytoplasm (31). Recently, NS2A was implicated in virus assembly, because specific mutations in this protein abolished production of infectious particles (16). In addition, it was proposed that NS3 also plays a role in assembly (14, 16, 19). Since NS2A and NS3 are both located in the sites of RNA synthesis and taking into account that both these proteins bind viral genomic RNA (21), it is tempting to speculate that NS2A and NS3 are responsible for genome encapsidation, while NS4B comes into play during the next step, budding, possibly by facilitating specific interaction of the nucleocapsid and envelope. NS4B could also facilitate particle formation indirectly, via its involvement in the formation of virus-induced membrane structures essential for the secretion pathways (14b, 20a).

Finally, the suggestive nature of these sequencing data should be recognized. The evolutionary, enzymological, and molecular aspects discussed above require further investigation. The process of accumulation of viral RNAs in infected cells needs to be precisely quantitated using modern approaches such as real-time PCR. This could establish the exact number of rounds of RNA synthesis, confirm or correct our estimate of YFpol fidelity, and add to a better understanding of the molecular mechanisms of flavivirus RNA synthesis. The high fidelity of YFpol requires confirmation by other available methods, including using wild-type YF. The hypothetical role of NS4B (as well as the elements in prM-E) in particle assembly requires direct demonstration.

This work was funded by Aventis Pasteur, Lyon, France.

REFERENCES

1. **Allison, S. L., K. Stadler, C. W. Mandl, C. Kunz, and F. H. Heinz.** 1995. Synthesis and secretion of recombinant tick-borne encephalitis virus protein E in soluble and particulate form. J. Virol. 69:5816–5820.

1a. **Brinton, M. A.** 1986. Replication of flaviviruses, p. 327. *In* S. Schlesinger and M. J. Schlesinger (ed.), The Togaviridae and Flaviviridae. Plenum Press, New York, N.Y.

2. **Burke, D. S., and T. P. Monath.** 2001. Flaviviruses, p. 1043–1126. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), Fields virology, 4th ed. Lippincott Williams and Wilkins, Philadelphia, Pa.

3. **Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice.** 1990. Flavivirus genome organization, expression, and replication. Annu. Rev. Microbiol. 44:649–688.

4. **Chambers, T. J., A. Nestorowicz, P. W. Mason, and C. M. Rice.** 1999. Yellow fever/Japanese encephalitis chimeric virus: construction and biological properties. J. Virol. 73:3095–3101.

5. **Chumakov, K. M., L. B. Powers, K. E. Noonan, I. B. Roninson, and I. S. Levenbook.** 1991. Correlation between amount of virus with altered nucleotide sequence and the monkey test for acceptability of oral poliovirus vaccine. Proc. Natl. Acad. Sci. USA 88:199–203.

6. **Domingo, E., C. Escarmis, N. Sevilla, A. Moya, S. F. Elena, J. Quer, I. S. Novella, and J. J. Holland.** 1996. Basic concepts in RNA virus evolution. FASEB J. 10:859–864.

7. **Drake, J. W., B. Charlesworth, D. Charlesworth, and J. F. Crow.** 1998. Rates of spontaneous mutation. Genetics 148:1667–1686.

8. **Elena, S. F., R. Mirales, J. M. Cuevas, P. E. Turner, and A. Moya.** 2000. The two faces of mutation: extinction and adaptation in RNA viruses. IUBMB Life 49:5–9.

9. **Elena, S. F., and A. Moya.** 1999. Rate of deleterious mutation and the distribution of its effects on fitness in vesicular stomatitis virus. J. Evol. Biol. 12:1078–1088.

10. **Guirakhoo, F., J. Arroyo, K. V. Pugachev, C. Miller, Z.-X. Zhang, R. Weltzin, K. Georgakopoulos, J. Catalan, S. Ocran, K. Soike, M. Ratterree, and T. P. Monath.** 2001. Construction, safety, and immunogenicity in nonhuman primates of a chimeric yellow fever-dengue virus tetravalent vaccine. J. Virol. 75:7290–7304.

11. **Guirakhoo, F., K. Pugachev, J. Arroyo, C. Miller, Z.-X. Zhang, R. Weltzin, K. Georgakopoulos, J. Catalan, S. Ocran, K. Draper, and T. P. Monath.** 2002. Viremia and immunogenicity in nonhuman primates of a tetravalent yellow fever-dengue chimeric vaccine: genetic reconstructions, dose adjustment, and antibody responses against wild-type dengue virus isolates. Virology 298:146–159.

12. **Hahn, C. S., J. M. Dalrymple, J. H. Strauss, and C. M. Rice.** 1987. Comparison of the virulent Asibi strain of yellow fever virus with the 17D vaccine strain derived from it. Proc. Natl. Acad. Sci. USA 84:2019–2023.

13. **Huang, J., L. G. Brieba, and R. Sousa.** 2000. Misincorporation by wild-type and mutant T7 RNA polymerases: identification of interactions that reduce misincorporation rates by stabilizing the catalytically incompetent open conformation. Biochemistry 39:11571–11580.

14. **Khromykh, A. A., P. L. Sedlak, and E. G. Westaway.** 2000. *cis*- and *trans*-acting elements in flavivirus RNA replication. J. Virol. 74:3253–3263.

14a. **Kofler, R. M., F. X. Heinz, and C. D. Mandl.** 2002. Capsid protein C of tick-borne encephalitis virus tolerates large internal deletions and is a favorable target for attenuation of virulence. J. Virol. 76:3534–3543.

14b. **Konan, K. V., T. H. Giddings, Jr., M. Ikeda, K. Li, S. M. Lemon, and K. Kirkegaard.** 2003. Nonstructural protein precursor NS4A/B from hepatitis C virus alters function and ultrastructure of host secretory apparatus. J. Virol. 77:7843–7855.

15. **Kuhn, R. J., W. Zhang, M. G. Rossmann, S. V. Pletnev, J. Corver, E. Lenches, C. T. Jones, S. Mukhopadhyay, P. R. Chipman, E. G. Strauss, T. S. Baker, and J. H. Strauss.** 2002. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell 108:717–725.

16. **Kummerer, B. M., and C. M. Rice.** 2002. Mutations in the yellow fever virus nonstructural protein NS2A selectively block production of infectious particles. J. Virol. 76:4773–4784.

17. **Lin, C., S. M. Amberg, T. J. Chambers, and C. M. Rice.** 1993. Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 proteinase is a prerequisite for processing at the downstream 4A/4B signalase site. J. Virol. 67:2327–2335.

18. **Lindenbach, B. D., and C. M. Rice.** 2001. Flaviviridae: the viruses and their replication, p. 991–1041. *In* D. M. Knipe, P. M. Howley, D. E. Griffin, R. A. Lamb, M. A. Martin, B. Roizman, and S. E. Straus (ed.), Fields virology, 4th ed. Lippincott Williams and Wilkins, Philadelphia, Pa.

19. **Liu, W. J., P. L. Sedlak, N. Kondratieva, and A. A. Khromykh.** 2002. Complementation analysis of the flavivirus Kunjin NS3 and NS5 proteins defines the minimal regions essential for formation of a replication complex and shows a requirement of NS3 in *cis* for virus assembly. J. Virol. 76:10766–10775.

20. **Lorenz, I. C., S. L. Allison, F. X. Heinz, and A. Helenius.** 2002. Folding and dimerization of tick-borne encephalitis virus envelope proteins prM and E in the endoplasmic reticulum. J. Virol. 76:5480–5491.

20a. **Lundin, M., M. Monne, A. Widell, G. Von Heijne, and M. A. Persson.** 2003. Topology of the membrane-associated hepatitis C virus protein NS4B. J. Virol. 77:5428–5438.

21. **Mackenzie, J. M., A. A. Khromykh, M. K. Jones, and E. G. Westaway.** 1998. Subcellular localization and some biochemical properties of the flavivirus Kunjin nonstructural proteins NS2A and NS4A. Virology 245:203–215.

22. **Mandl, C. W., H. Kroschewski, S. L. Allison, R. Kofler, H. Holzmann, T. Meixner, and F. H. Heinz.** 2001. Adaptation of tick-borne encephalitis virus to BHK-21 cells results in the formation of multiple heparan sulfate binding sites in the envelope protein and attenuation in vivo. J. Virol. 75:5627–5637.

23. **Meyerhans, A., and J.-P. Vartanian.** 1999. The fidelity of cellular and viral polymerases and its manipulation for hypermutagenesis, p. 87–114. *In* E. Domingo, R. Webster, and J. Holland (ed.), Origin and evolution of viruses. Academic Press, New York, N.Y.

24. **Monath, T. P., K. Soike, I. Levenbook, Z.-X. Zhang, J. Arroyo, S. Delagrave, G. Myers, A. D. T. Barrett, R. E. Shope, M. Ratterree, T. J. Chambers, and F. Guirakhoo.** 1999. Recombinant, chimaeric live, attenuated vaccine (ChimeriVax) incorporating the envelope genes of Japanese encephalitis (SA14–14–2) virus and the capsid and nonstructural genes of yellow fever (17D) virus is safe, immunogenic and protective in non-human primates. Vaccine 17:1869–1882.

25. **Monath, T. P., J. Arroyo, C. Miller, and F. Guirakhoo.** 2001. West Nile virus vaccine. Curr. Drug Trends Infect. Disorders 1:37–50.

26. **Monath, T. P., K. McCarthy, P. Bedford, C. T. Johnson, R. Nichols, S. Yoksan, R. Marchesani, M. Knauber, K. H. Wells, J. Arroyo, and F. Guirakhoo.** 2002. Clinical proof of principle for ChimeriVax: recombinant live, attenuated vaccines against flavivirus infections. Vaccine 20:1004–1018.

27. **Mutebi, J.-P., H. Wang, L. Li, J. E. Bryant, and A. D. T. Barrett.** 2001. Phylogenetic and evolutionary relationships among yellow fever virus isolates in Africa. J. Virol. 75:6999–7008.

27a. **Pugachev, K. V., S. W. Ocran, F. Guirakhoo, D. Furby, and T. P. Monath.** 2002. Heterogeneous nature of the genome of the ARILVAX yellow fever 17D vaccine revealed by consensus sequencing. Vaccine 20:996–999.

28. **Ramirez, B.-C., P. Barbier, K. Seron, A.-L. Haenni, and F. Bernardi.** 1995. Molecular mechanisms of point mutations in RNA viruses, p. 105–118. *In* A. Gibbs, C. H. Calisher, and F. García-Arenal (ed.), Molecular basis of virus evolution. Cambridge University Press, Cambridge, England.

29. **Rey, F. A., F. X. Heinz, C. Mandl, C. Kunz, and S. C. Harrison.** 1995. The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. Nature 375:291–298.

30. **Westaway, E. G.** 1980. Replication of flaviviruses, p. 531–581. *In* The togaviruses: biology, structure, replication. Academic Press, New York, N.Y.

31. **Westaway, E. G., A. A. Khromykh, M. T. Kenney, J. M. Mackenzie, and M. K. Jones.** 1997. Proteins C and NS4B of the flavivirus Kunjin translocate independently into the nucleus. Virology 234:31–41.

32. **Westaway, E. G., J. M. Mackenzie, M. T. Kenney, M. K. Jones, and A. A. Khromykh.** 1997. Ultrastructure of Kunjin virus-infected cells: colocalization of NS1 and NS3 with double-stranded RNA, and of NS2B with NS3, in virus-induced membrane structures. J. Virol. 71:6650–6661.

33. **Xie, H., A. R. Cass, and A. D. T. Barrett.** 1998. Yellow fever 17D vaccine virus isolated from healthy vaccines accumulates very few mutations. Virus Res. 55:93–99.